

SYSTEMATIC REVIEW

FUNCTIONAL PERFORMANCE TESTING OF THE HIP IN ATHLETES: A SYSTEMATIC REVIEW FOR RELIABILITY AND VALIDITY

Benjamin R. Kivlan, PT, SCS, OCS^{1,2}RobRoy L. Martin, PhD, ATC^{1,3}

ABSTRACT

Purpose/Background: The purpose of this study was to systematically review the literature for functional performance tests with evidence of reliability and validity that could be used for a young, athletic population with hip dysfunction.

Methods: A search of PubMed and SPORTDiscus databases were performed to identify movement, balance, hop/jump, or agility functional performance tests from the current peer-reviewed literature used to assess function of the hip in young, athletic subjects.

Results: The single-leg stance, deep squat, single-leg squat, and star excursion balance tests (SEBT) demonstrated evidence of validity and normative data for score interpretation. The single-leg stance test and SEBT have evidence of validity with association to hip abductor function. The deep squat test demonstrated evidence as a functional performance test for evaluating femoroacetabular impingement. Hop/Jump tests and agility tests have no reported evidence of reliability or validity in a population of subjects with hip pathology.

Conclusions: Use of functional performance tests in the assessment of hip dysfunction has not been well established in the current literature. Diminished squat depth and provocation of pain during the single-leg balance test have been associated with patients diagnosed with FAI and gluteal tendinopathy, respectively. The SEBT and single-leg squat tests provided evidence of convergent validity through an analysis of kinematics and muscle function in normal subjects. Reliability of functional performance tests have not been established on patients with hip dysfunction. Further study is needed to establish reliability and validity of functional performance tests that can be used in a young, athletic population with hip dysfunction.

Level of Evidence: 2b (Systematic Review of Literature)

Key Words: Functional Testing, hip, reliability, validity

CORRESPONDING AUTHOR

Benjamin R. Kivlan, PT, SCS, OCS
Rangos School of Health Sciences
Duquesne University
Pittsburgh, PA 15282
724-799-4065
bkivlan@zoominternet.net

¹ John G. Rangos Sr. School of Health Sciences, Duquesne University, Pittsburgh, PA

² Tri-State Physical Therapy, Seven Fields, PA

³ UPMC Center for Sports Medicine, Pittsburgh, PA

INTRODUCTION

Over the past decade, physical therapists have adopted the International Classification of Functioning, Disability, and Health (ICF) model¹ to evaluate and treat patients of all levels of health and disability. The ICF model encourages assessment of patients, including athletes, within the context of their function. Therefore, an athlete, or a subject that is regularly participating in an organized sport activity or training, should be evaluated with consideration of the physical demands relative to their sport-related activities. Current evaluation procedures for the hip including range of motion (ROM), strength, and special tests are intended to identify a specific pathology or impairment,² not necessarily to identify dysfunction. Dysfunction can be defined as pain, asymmetry, or injury that impairs normal movement and performance of a functional activity. Clinicians need to understand which sport-specific functional performance tests can best be used to identify and treat hip dysfunction of an athlete.³

Functional performance tests require the integration of multiple body regions and systems to execute movement patterns and therefore may have an advantage over more traditional clinical measures. Components of ROM, flexibility, muscular strength, endurance, coordination, balance, and motor control of multiple regions can be assessed simultaneously by observing the movement patterns in which the athlete normally functions.^{4,5,6} Functional performance tests have been commonly used to identify impairments related to ankle⁷⁻¹⁵ or knee injuries^{16,17,18-21} and determine the readiness of an athlete to return to sports after injury.²² Similar information has not been established for the hip. Although athletes with hip dysfunction are commonly encountered in a sports medicine rehabilitation setting, currently it is unclear which functional performance tests are most appropriate to use in this population.

To help clinicians employ appropriate functional performance tests in the evaluation of hip dysfunction, the evidence for reliability and validity of the functional performance tests should be considered. Reliability describes how well the test can be reproduced under the same conditions. Validity describes how well a test measures what it is intended to measure. The population in which the evidence for reliability

and validity of a functional performance test is established is also an important consideration. Evidence of reliability and validity should be established among a sample of subjects that are similar to the population of patients for which the test is to be used.²⁴ Clinicians need to determine which functional performance tests are valid and reliable for use during evaluation of athletes with hip dysfunction. The purpose of this systematic review was to examine current evidence for reliability and validity of functional performance tests used in a young, athletic population with hip dysfunction.

METHODS

Search Strategy

An electronic search of PubMed and SPORTDiscus databases were performed for the purpose of identifying peer-reviewed articles that utilized functional performance tests to assess function of the hip joint and related structures. The following key words were used in combination for the search of the respective databases: “functional” AND “test”, OR “measure”, OR “assessment”, OR “screen”, AND “lower extremity” OR “hip”. The primary author reviewed the abstracts of the articles identified from the database searches and duplicate studies were removed. The full text articles describing a functional performance test related to the hip joint were retrieved for data extraction and the references reviewed for additional articles of interest. In addition, the individual names of the identified functional performance tests including the deep squat, single-leg squat, single-leg balance, Star Excursion Balance Test, Balance Error Scoring System, Functional Movement Screen™, single leg hop, triple hop, side hop, timed hop, medial hop, lateral hop, figure 8 hop, cross-over hop, square hop, agility T-test, modified agility T-test, and reactive agility test were searched independently to provide a comprehensive list of functional performance tests used for patients with hip dysfunction.

Study Selection Criteria

Studies were included if the following criteria were met: 1) written in English 2) published in a peer-reviewed journal from 1990-present, and 3) described the use of a functional performance test in the assessment of ROM, strength, balance, postural control, or athletic performance of the lower extremity. Excluded

studies were those that described patient self-report measures, targeted an elderly population (>60 years of age), or if the functional performance test was performed on subjects with neurological involvement caused by cerebral palsy, a cardiovascular accident, or head injury.

Data Extraction

Grouping of Functional Performance Tests

Functional performance tests were grouped into one of four categories: 1) movement 2) balance, 3) hop/jump, or 4) agility. A movement test was defined as a quantitative or qualitative measure of the subject's ability to perform a specific motor pattern with control, precision, and symmetry. Tests of balance measured the patient's ability to maintain balance and postural control under different conditions. Hop/jump tests assessed the quality and/or quantity of tasks related to propulsion and/or absorbing impact. Agility tests assessed the subject's ability to run, cut, pivot, and/or change direction through a predetermined course.

Evidence of Reliability

Statistical evidence for test-retest, intra-rater, or inter-rater reliability was recorded as it was originally reported in the identified articles. This included intra-class correlation coefficients (ICC) for interval data or a kappa statistic for ordinal and normative data. An ICC or Kappa statistic can range from 0.0 to 1.0. An ICC or Kappa statistic greater than or equal to .75 are considered excellent, .40 to .74 are moderate, and less than .40 are considered poor.²⁵ Test-retest reliability refers to the agreement of two separate trials of the same test by a single investigator.²⁶ Intra-rater reliability is the agreement of a singular testing session by single investigator. This is usually established through videotaped recordings of a singular testing session. Inter-rater reliability is the agreement of the same test by two separate investigators. Intra-rater reliability and test-retest reliability are similar and often referred interchangeably in the literature. There is, however, an important distinction between the intra-rater and test-retest reliability. Intra-rater reliability isolates intra-rater error as the same performance is evaluated on two separate occasions (i.e. a single evaluator watches the videotape and grades a single test performance two separate times).²⁶ Test-retest reliability also

accounts for intra-rater error in addition to the variability of performance on two separate testing sessions (i.e. a single evaluator grades the performance of subjects performing a functional performance test on two separate occasions).²⁶ Test-retest reliability more closely represents how functional performance tests may likely be employed in a clinical environment. The authors of the original articles reviewed for this analysis may not have intended to imply the terminology as defined above. Therefore, when reporting reliability the terminology consistent to that reported in the original article was used.

Evidence of validity

Evidence of validity of a functional performance test can be established by demonstrating its relationship to subjects with a known dysfunction of the hip joint. Such a relationship may be expressed as a value of sensitivity or specificity in detecting the presence of dysfunction.²⁵ Validity may also be established by a relationship, expressed statistically by a correlation coefficient, to other variables of hip function.²⁵ For instance, a relationship of a functional performance test to ROM or strength values of the hip joint may offer evidence of validity to a functional performance test.

Evidence of score interpretation

Evidence for score interpretation was also extracted when available. This included normative values as well as the smallest detectable difference (SDD) or the minimal detectable change (MDC) as reported in the original article. The SDD and MDC are often described interchangeably and represent the change or difference between test scores that distinguish error from true changes in the measurement.²³

RESULTS

The search results consisted of 18 articles describing movement tests, 24 articles describing balance tests, 26 articles describing hop/jump tests, and 6 articles describing agility tests as shown in Figure 1. Review of the articles revealed four functional performance tests that demonstrated evidence of validity. This included two movement tests (the deep squat test²⁷ and single-leg squat test^{28,29}) and two balance tests (single leg stance test³⁰ and Star Excursion Balance Test).^{31,32,33} The deep squat was the only functional

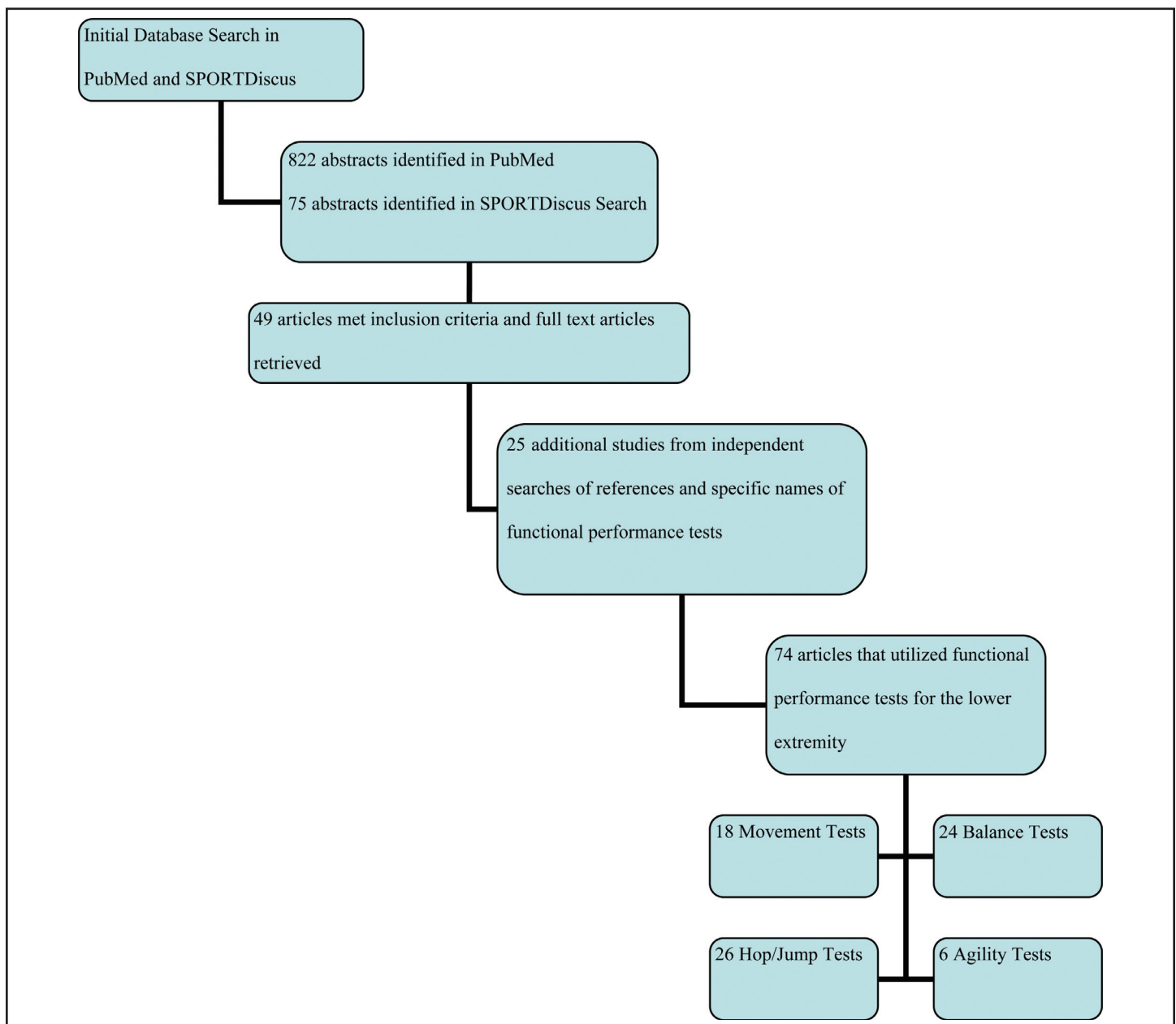


Figure 1. Results of Literature Search for Functional Performance Tests.

performance test that was performed on subjects with hip dysfunction.²⁷ During the deep squat test, patients with femoroacetabular impingement (FAI) demonstrated less squat depth and altered lumbo-pelvic kinematics compared to healthy controls.²⁷ The single leg squat, single-leg stance, and the Star Excursion Balance Test (SEBT) were functional performance tests that demonstrated evidence of validity through a relationship to hip abductor muscle function. Subjects graded as “poor” on the single leg squat test exhibited weaker and slower muscle

activation of the hip abductors than those graded as “good”.^{28,29} Provocation of pain during 30-second single-leg stance had high sensitivity (100%) and specificity (97.3%) in detecting tendinopathy of the gluteus medius and minimus.³⁰ For the SEBT, posterior-medial and posterior-lateral reach distances of the SEBT have been correlated to hip abduction and extension strength ($r = .48 - .51$)³¹, and the medial reach of the SEBT elicited activation of the gluteus medius at 49% of maximal volitional isometric contraction.³³ The SEBT also demonstrated a relationship

Table 1. Summary of Functional Performance Tests with Evidence of Validity to Hip Function.			
Test	Category	Relationship to Hip Function	Interpretation of Test Results
Deep Squat	Movement	<ul style="list-style-type: none"> Patients with femoroacetabular impingement demonstrated less squat depth and altered lumbo-pelvic kinematics²⁷ 	<ul style="list-style-type: none"> Patients with Femoroacetabular Impingement Mean Peak Squat Depth = 41% of leg length Normal Mean Peak Squat Depth = 32% of leg length²⁷
Single-Leg Squat	Movement	<ul style="list-style-type: none"> Subjects graded as “poor” on the single leg squat test exhibit weaker and slower muscle activation of the hip abductors than those graded as “good”.²⁹ 	<ul style="list-style-type: none"> Subjects were ordinally graded (Good, Fair, Poor) on ability to maintain balance, postural control, and lower body alignment during 5 repetitions of a single leg squat to 60°.²⁹
Single-Leg Stance	Balance	<ul style="list-style-type: none"> Provocation of pain during 30-second single-leg stance has shown sensitivity (100%) and specificity (97.3%) in detecting tendinopathy of the gluteus medius and minimus.³⁰ 	<ul style="list-style-type: none"> Positive Test = increase of pain within 30 seconds of single leg stance³⁰ Normal function of the hip abductors maintains the pelvis nearly perpendicular to the femur in a single leg stance position.⁴⁶ Normal = 30 seconds of single leg stance without pain³⁰
Star Excursion Balance Test	Balance	<ul style="list-style-type: none"> Hip flexion range of motion explained 86-92% of SEBT reach distances³² Hip abduction and extension strength have a moderate correlation ($r=.48 - .51$) to posterior-medial and posterior-lateral reach distances³¹ Elicits activation of the gluteus medius at 49% of maximal volitional isometric contraction during medial reach.³³ 	<ul style="list-style-type: none"> Anterior reach difference ≥ 4 cm is 2.5 times risk for injury.⁴⁷ A composite score standardized to leg length $<94\%$ is 6 times more likely for injury⁴⁷

to hip ROM. Hip flexion ROM was shown to explain a high percentage of variance (92-95%) in SEBT scores.³² Table 1 summarizes the evidence of validity for the use of functional performance tests in subjects with hip dysfunction. There was no evidence of validity for hop/jump or agility tests for patients with hip dysfunction.

None of the functional performance tests demonstrated evidence of reliability in a population of young, athletic patients with hip dysfunction. However, evidence of reliability for functional performance tests in healthy subjects with normative values for score interpretation was identified in 2 movement, 4 balance, 11 hop/jump, and 3 agility tests (Table 2). Ten of these functional performance tests also had established SDD or MDC values on healthy subjects. This information is summarized in Table 2. The authors from a majority of the articles reported moderate to excellent reliability (.61-1.00) of the functional performance tests in groups of healthy subjects. The single leg squat, single leg balance, and star excursion balance test, however, had conflicting evidence of reliability with some reports suggesting poorer reliability (.21-.58) of these specific functional performance tests.

DISCUSSION

The primary purpose of this study was to systematically review the literature for evidence of reliability and validity of functional performance tests used with a young, athletic population with hip dysfunction. The single-leg stance test, deep squat test, single-leg squat test, and SEBT demonstrated evidence of validity to be used in a population of patients with suspected hip dysfunction. The evidence for validity suggests that gluteal tendinopathy and function of the hip abductors may be assessed with the single-leg stance test,³⁰ single-leg squat test,^{28,29} and SEBT.^{32,33} The deep squat test demonstrated evidence of validity as a functional performance test for evaluating patients with suspected FAI.²⁷ Additionally, there were 20 tests that had evidence of reliability in healthy subjects with normative data provided to aide in score interpretation. Clinicians may use this normative data to identify impairments of patients with hip dysfunction that score outside the normal range of healthy subjects on the functional performance tests.

The authors of the current systematic review aimed to identify existing functional performance tests that demonstrated evidence for validity in a population of young, athletic patients with hip dysfunction. Only

Table 2. Evidence of Reliability and Normative Values for Functional Performance Tests in Physically Active Subjects.

Test	Category	Average age of subjects	Evidence of Reliability	Normative values for Healthy Subjects
Single Leg Squat Test	Movement	24 ± 5 years ²⁹	.61 - .80 (intra-rater) ²⁹	“Good” to “Excellent” alignment of hip flexion (<65°), hip adduction (<10°) and knee valgus (<10°) ²⁹
Functional Movement Screen	Movement	22 ± 4 years ³⁷	.97 (inter-rater) ³⁷	15-16 composite score ³⁷
Single Leg Balance Test (Trendelenberg Test)	Balance	33± 11 years (males) 28± 8 years(females) ⁴⁶	.58 (intra-rater) ⁴⁶ MCD >4° of pelvic on femoral angle ⁴⁶	Normal pelvic on femoral angle = 83° (range:76-94°) ⁴⁶
Single Leg Balance (10 second Balance Test with eyes closed)	Balance	21± 2 years (males) 28± 8 years (females) ⁴⁸	.21 (test-retest) 1.00 (inter-rater) ⁴⁸	Maintain postural control and balance on a single limb with eyes closed for 10 seconds ⁴⁸
Balance Error Scoring System	Balance	21± 2 years (males) 28± 8 years (females) ⁴⁸	.74 (test-retest) ⁴⁸ SDD= 9.4 errors (inter-tester) 7.3 errors (intra-tester) ⁴⁹	9.1 errors - female gymnasts 14.1 errors - female basketball players ⁵⁰ 5-7 errors – male baseball players ⁵¹
Star Excursion Balance Test	Balance	22 ± 4 years (males) 23± 3 years (females) ⁵²	.84-.92 (test-retest) ⁵² .35-.93 (inter-rater) ⁵³ SDD= 6-8% of the subject’s limb length ⁵²	Anterior reach difference < 4 cm >94% of composite score ⁴⁷
Single Leg Hop	Hop/Jump	23 ± 3 years (males) 22 ± 4 years (females) ⁴³ ; 25± 4 years ⁵⁴	.80- .96 (test-retest) ^{43,54} SDD= 22% of the subject’s limb length ⁴³	Females=157% of leg length Males=189% of leg length ⁴³
Triple Hop	Hop/Jump	23 ± 3 years (males) 22 ± 4 years (females) ⁴³ ; 25± 4 years ⁵⁴	.80-.95 (test-retest) ^{43,54} SDD: Females= 48% of leg length Males= 64% of leg length ⁴³	Females= 505% of leg length Males=585% of leg length ⁴³
Timed Hop Test	Hop/Jump	23 ± 3 years (males) 22 ± 4 years (females) ⁴³ ; 25± 4 years ⁵⁴	.60 - .84 (test-retest) ^{43,54} SDD: Females= .21 seconds Males= .23 seconds ⁴³	Females= 2.06 seconds Males=1.76 seconds LSI > 99% ⁴³
One-legged Hop Test (Medial and Lateral)	Hop/ Jump	Male hockey players 20± 3 years ⁵⁵	.87-.95 (test-retest) ⁵⁵	Medial=157 cm and Lateral=160 cm for male hockey players 2% difference between medial and lateral distances ⁵⁵
Side Hop Test (Timed Test)	Hop/Jump	female athletes 22± 2 years ²⁰ ; college students 20± 2 years ⁵⁶	.48-.84 (test-retest) ^{20,56} MDC=5.82 seconds ⁵⁶	Females=8.20 seconds Males=7.36 seconds ¹⁹
Side Hop Test (Number in 30 seconds)	Hop/Jump	Male and female 28± 4 years ¹⁸	.87 – 93 (test-retest) ¹⁸	50 hops ¹⁸
Figure 8 Hop Test	Hop/Jump	female athletes 22± 2 years ²⁰ ; college students 20± 2 years ⁵⁶	.85-.99 (test-retest) ^{20,56} MDC=4.59 seconds ⁵⁶	Females=12.47 seconds Males=11.36 seconds LSI ≤ .81 seconds ¹⁹
The Cross-Over Hop Test (for time)	Hop/Jump	college students 20± 2 years ⁵⁶	.96 (test-retest) MDC=1.03 seconds ⁵⁶	2.7 -2.8 seconds ⁵⁶
The Cross-Over Hop Test (for distance)	Hop/Jump	23 ± 3 years (males) 22 ± 4 years (females) ⁴³ ; 25± 4 years ⁵⁴	.86-.96 (test-retest) ^{43,54} SDD=55-59% of the subject’s leg length ⁵⁴	Females= 480% of leg length Males=555% of leg length ⁵⁴
Square Hop Test (for time)	Hop/Jump	college students 20± 2 years ⁵⁶	.90 (test-retest) MDC=3.88 seconds ⁵⁶	15 – 16 seconds ⁵⁶

Table 2. Evidence of Reliability and Normative Values for Functional Performance Tests in Physically Active Subjects. (continued)

Square Hop Test (for number of jumps)	Hop/Jump	Male and female 28± 4 years ¹⁸	.58 - .85 (test-retest) ¹⁸	62 jumps ¹⁸
Agility T-Test	Agility	23 ± 3 years (males) 22 ± 4 years (females) ⁴³	.82-.96 (test-retest) SDD: Male = .48 seconds Female = .58 seconds ⁴³	Male = 10.7 seconds Female = 13.0 seconds ⁴³
Modified Agility Test	Agility	Male and female 21± 4 years ⁴⁴	.83 (test-retest) ⁴⁴	9.59 seconds ⁴⁴
Reactive Agility Test	Agility	Male Australian Football players 17±1 years ⁴⁵	.91 (test-retest) ⁴⁵	2.42 – 2.96 seconds ⁴⁵

SDD=smallest detectable difference; MCD=minimal detectable change; LSI=limb symmetry index.

the single-leg stance test and deep squat test had evidence of validity in patients with confirmed hip dysfunction. The single-leg stance test was performed on subjects with greater than 4 months of lateral hip pain. Provocation of pain during 30-seconds of single leg stance had high sensitivity (100%) and specificity (97.3%) in detecting tendinopathy of the gluteus medius and minimus confirmed by magnetic resonance imaging (MRI).³⁰ Based on the current evidence, the single-leg stance test has clinical value in ruling out other potential sources of lateral hip pain including lumbosacral, sacroiliac, or intra-articular pathology from gluteal tendinopathy, otherwise known as greater trochanteric pain syndrome (GTPS).³⁰ The deep squat test was performed on subjects with radiologically confirmed FAI. The maximal squat depth in subjects with FAI (41 % of leg length) was significantly less when compared to healthy controls (32% of leg length).²⁷ Clinicians may test maximum squat depth in patients with suspected FAI to help confirm a diagnosis of FAI. Further studies are needed to determine how the single-leg stance and deep squat tests may compliment current clinical exam procedures to identify the presence of specific hip dysfunction.

While the single-leg stance and deep squat tests provided evidence of validity in subjects with hip dysfunction, the SEBT and single-leg squat test provided evidence of validity through an analysis of kinematics and muscle function in normal subjects. Three studies have related SEBT performance to kinematic and muscle function variables of the hip joint. Hip flexion range of motion was shown to explain a high percentage of variance (92-95%) in SEBT perfor-

mance.³² Electromyographic study of the gluteus medius demonstrated the medial reach of the SEBT elicited the gluteus medius at 49% of maximal volitional isometric contraction.³³ Hip abduction and extension strength also demonstrated a moderate correlation ($r = .48 - .51$) to posterior-medial and posterior-lateral reach distances of the SEBT.³¹ The moderate correlation implies that gluteal muscle strength only partially accounts for the variance of SEBT scores. The single-leg squat also demonstrated a relationship to hip abductor muscle function.²⁹ However, the strength of this relationship has been disputed. DiMattia et al.²⁸ reported poor association ($r = .21$) of the single-leg squat to hip abductor strength. The SEBT and the single-leg squat test have not been studied on a population of patients with hip dysfunction, but may have value to help clinicians screen for ROM and muscle strength impairments.²⁸⁻³³ ROM and strength deficits are commonly associated with hip pathology including FAI,³⁴ osteoarthritis,³⁵ or GTPS.³⁶ A positive finding or asymmetry on the SEBT or single-leg squat test may lead the clinician to perform goniometry or dynamometry to further objectify ROM and strength deficits observed during the functional performance tests.

The reliability of a test is important to be able to confidently interpret the results. There were 2 movement, 4 balance, 11 hop/jump, and 3 agility tests with evidence of reliability in a young, healthy population. The Functional Movement Screen™ (FMS) is a series of seven individual movement tests that have been reliable in screening and evaluating athletes.³⁷ Each test is graded on an ordinal scale based on the ability of the subject to perform specific motor func-

tions.³⁷ The FMS™ is designed to be a comprehensive cross section of functional movement and has been used to predict an athlete's risk for non-specific injury.³⁸ The FMS™ is an intriguing tool for patients with varied hip dysfunction as it tests multiple movement patterns that require different components of hip ROM, strength, and trunk control. Such tests may elicit familiar symptoms or indicate impairments related to FAI, labral tears, osteoarthritis, or GTPS. Clinicians may use normative data established for the FMS™ as a guide to identify abnormal findings on the FMS™ for patients with hip-related dysfunction.³⁷ Further study is needed to determine if the FMS™ is able to accurately predict hip-specific injuries.

Similar to the FMS™, hop/jump tests also demonstrated good to excellent reliability in normal subjects. Hop tests have also shown ability to discriminate injured from uninjured lower extremities, particularly in the assessment of ankle instability and post-operatively following ACL reconstruction.³⁹⁻⁴² Researchers have established normative, gender-specific values for hop tests^{19,43} in young, athletic subjects. These values may serve as benchmarks that may be helpful in interpreting an "abnormal" score for a subject with hip-related dysfunction. Field agility tests have demonstrated evidence of good reliability,^{19,43-45} but have not been able to discriminate injured versus uninjured limbs in the same manner as hop tests. This is likely because agility tests require bipedal movement. However, agility tests may have value in an athletic population as the tests may more closely mimic the dynamic requirements of sport activity. The reliability of hop/jump and agility tests measures have not been established on patients with hip dysfunction. It remains unclear how patients with hip dysfunction perform on these tests without further study. For patients with unilateral hip symptoms, hop tests may be used in comparison of the uninvolved side. Interpretation of agility test results is limited to a comparison of scores established on healthy subjects. Whether jump/hop tests or agility tests can be used to discriminate subjects with hip-related dysfunction remains unknown.

There are limitations of this systematic review that should be acknowledged. First, the authors established very specific inclusion/exclusion criteria for selection of functional performance tests included in

this review. This included only exploring functional performance tests for a young and athletic population. Many tests were excluded because the studies were performed on elderly patients, or subjects with various neurologic or debilitating co-morbidities. Therefore, a number of articles that examined functional performance tests did not fit the inclusion criteria. It is possible some of these functional performance tests may have value in a population of subjects with musculoskeletal dysfunction. Given the volume of articles reviewed, it is also possible some functional performance tests were not identified. Significant variations in the descriptions of functional performance tests used by different authors were common. This may alter interpretation of the values attained for a specific functional performance test.

In conclusion, only the deep squat and single-leg stance test demonstrated evidence of validity in a population of patients with hip-related dysfunction. Specifically, diminished squat depth and provocation of pain during the single-leg balance test may be an indication for FAI and gluteal tendinopathy, respectively. The SEBT and single-leg squat test provided evidence of validity through an analysis of kinematics and muscle function in healthy subjects. There were 20 functional performance tests, including the FMS™, with evidence of reliability and normative data to help in score interpretation. None of the articles provided evidence of reliability in a group of subjects with hip-related dysfunction. Without established reliability for these functional performance tests it limits the ability of the clinician to confidently interpret test results and utilize the tests to measure patient progress. The results of the systematic review demonstrated few functional performance tests that have established validity and reliability to compliment traditional clinical exam procedures for patients with hip dysfunction. Further study is needed to establish the reliability and validity of existing functional performance tests or explore new, relevant functional performance tests to be used in a young, athletic population with hip dysfunction.

REFERENCES

1. World Health Organization International Classification of Functioning, Disability, and Health. Geneva, Switzerland: 2001.

2. Martin HD, Kelly BT, Leunig M, Philippon MJ, Clohisy JC, Martin RL, Sekiya JK, Pietrobon R, Mohtadi NG, Sampson TG, Safran MR. The pattern and technique in the clinical evaluation of the adult hip: the common physical examination tests of hip specialists. *Arthroscopy*. 2010 Feb;26(2):161-72. PubMed PMID: 20141979.
3. Reiman MP, Manske RC. *Functional testing in human performance*. Champaign, IL: Human Kinetics; 2009:310.
4. Okada T, Huxel KC, Nesser TW. Relationship between core stability, functional movement, and performance. *J Strength Cond Res*. 2011;25(1):252-261. doi: 10.1519/JSC.0b013e3181b22b3e.
5. Mills JD, Tauton J, Mills W. The effects of a 10 week training regimen on lumbo-pelvic stability and athletic performance in female athletes: a randomized controlled trial. *Phys Ther Sport*. 2005(6):60-66.
6. Cook G, Burton L, Hoogenboom B. Pre-participation screening: The use of fundamental movements as an assessment of function - part 1. *N Am J Sports Phys Ther*. 2006;1(2):62-72.
7. Arnold BL, De La Motte S, Linens S, Ross SE. Ankle instability is associated with balance impairments: A meta-analysis. *Med Sci Sports Exerc*. 2009;41(5):1048-1062. doi: 10.1249/MSS.0b013e318192d044.
8. Gribble PA, Hertel J, Denegar CR. Chronic ankle instability and fatigue create proximal joint alterations during performance of the star excursion balance test. *Int J Sports Med*. 2007;28(3):236-242.
9. Gribble PA, Hertel J, Denegar CR, Buckley WE. The effects of fatigue and chronic ankle instability on dynamic postural control. *J Athl Train*. 2004;39(4):321-329.
10. Hale SA, Hertel J, Olmsted-Kramer LC. The effect of a 4-week comprehensive rehabilitation program on postural control and lower extremity function in individuals with chronic ankle instability. *J Orthop Sports Phys Ther*. 2007;37(6):303-311.
11. Hertel J, Braham RA, Hale SA, Olmsted-Kramer LC. Simplifying the star excursion balance test: Analyses of subjects with and without chronic ankle instability. *J Orthop Sports Phys Ther*. 2006;36(3):131-137.
12. Martinez-Ramirez A, Lecumberri P, Gomez M, Izquierdo M. Wavelet analysis based on time-frequency information discriminate chronic ankle instability. *Clin Biomech (Bristol, Avon)*. 2010;25(3):256-264. doi: 10.1016/j.clinbiomech.2009.11.004.
13. McLeod TC, Armstrong T, Miller M, Sauers JL. Balance improvements in female high school basketball players after a 6-week neuromuscular-training program. *J Sport Rehabil*. 2009;18(4):465-481.
14. Olmsted LC, Carcia CR, Hertel J, Shultz SJ. Efficacy of the star excursion balance tests in detecting reach deficits in subjects with chronic ankle instability. *J Athl Train*. 2002;37(4):501-506.
15. Docherty CL, Arnold BL, Gansneder BM, Hurwitz S, Gieck J. Functional-performance deficits in volunteers with functional ankle instability. *J Athl Train*. 2005;40(1):30-34.
16. Aminaka N, Gribble PA. Patellar taping, patellofemoral pain syndrome, lower extremity kinematics, and dynamic postural control. *J Athl Train*. 2008;43(1):21-28. doi: 10.4085/1062-6050-43.1.21.
17. Herrington L, Hatcher J, Hatcher A, McNicholas M. A comparison of star excursion balance test reach distances between ACL deficient patients and asymptomatic controls. *Knee*. 2009;16(2):149-152. doi: 10.1016/j.knee.2008.10.004.
18. Gustavsson A, Neeter C, Thomee P, et al. A test battery for evaluating hop performance in patients with an ACL injury and patients who have undergone ACL reconstruction. *Knee Surg Sports Traumatol Arthrosc*. 2006;14(8):778-788. doi: 10.1007/s00167-006-0045-6.
19. Itoh H, Kurosaka M, Yoshiya S, Ichihashi N, Mizuno K. Evaluation of functional deficits determined by four different hop tests in patients with anterior cruciate ligament deficiency. *Knee Surg Sports Traumatol Arthrosc*. 1998;6(4):241-245.
20. Ortiz A, Olson S, Libby CL, Kwon YH, Trudelle-Jackson E. Kinematic and kinetic reliability of two jumping and landing physical performance tasks in young adult women. *N Am J Sports Phys Ther*. 2007;2(2):104-112.
21. Hewett TE, Myer GD, Ford KR, Slauterbeck JR. Preparticipation physical examination using a box drop vertical jump test in young athletes: The effects of puberty and sex. *Clin J Sport Med*. 2006;16(4):298-304.
22. Ardern CL, Webster KE, Taylor NF, Feller JA. Return to the preinjury level of competitive sport after anterior cruciate ligament reconstruction surgery: Two-thirds of patients have not returned by 12 months after surgery. *Am J Sports Med*. 2011;39(3):538-543. doi: 10.1177/0363546510384798.
23. Stauffer ME, Taylor SD, Watson DJ, Peloso PM, Morrison A. Definition of nonresponse to analgesic treatment of arthritic pain: An analytical literature review of the smallest detectable difference, the minimal detectable change, and the minimal clinically important difference on the pain visual analog scale. *Int J Inflam*. 2011;2011:231926. doi: 10.4061/2011/231926.
24. Martin RL, Irrgang JJ. A survey of self-reported outcome instruments for the foot and ankle. *J Orthop Sports Phys Ther*. 2007;37(2):72-84.

-
25. Andresen EM. Criteria for assessing the tools of disability outcomes research. *Archives of Physical Medicine and Rehabilitation*. 2000; 81,12(2):S15-S20.
26. Holmefur M, Aarts P, Hoare B, Krumlinde-Sundholm L. Test-retest and alternate forms reliability of the assisting hand assessment. *J Rehabil Med*. 2009;41(11):886-91. PubMed PMID: 19841839.
27. Lamontagne M, Kennedy MJ, Beaulé PE. The effect of cam FAI on hip and pelvic motion during maximum squat. *Clin Orthop Relat Res*. 2009;467(3):645-650. doi: 10.1007/s11999-008-0620-x.
28. Dimattia MA, Livengood AL, Uhl TL, Mattacola CG, Malone TR. What are the validity of the single-leg squat test and its relationship to hip abduction strength. *J Sport Rehabil*. 2005;14(2): 108-123.
29. Crossley KM, Zhang WJ, Schache AG, Bryant A, Cowan SM. Performance on the single-leg squat task indicates hip abductor muscle function. *Am J Sports Med*. 2011;39(4):866-873. doi: 10.1177/0363546510395456.
30. Lequesne M, Mathieu P, Vuillemin-Bodaghi V, Bard H, Djian P. Gluteal tendinopathy in refractory greater trochanter pain syndrome: Diagnostic value of two clinical tests. *Arthritis Rheum*. 2008;59(2):241-246. doi: 10.1002/art.23354.
31. Hubbard TJ, Kramer LC, Denegar CR, Hertel J. Correlations among multiple measures of functional and mechanical instability in subjects with chronic ankle instability. *J Athl Train*. 2007;42(3):361-366.
32. Robinson R, Gribble P. Kinematic predictors of performance on the star excursion balance test. *J Sport Rehabil*. 2008;17(4):347-357.
33. Norris B, Trudelle-Jackson E. Hip- and thigh-muscle activation during the star excursion balance test. *J Sport Rehabil*. 2011;20(4):428-441.
34. Beck M, Kalhor M, Leunig M, Ganz R. Hip morphology influences the pattern of damage to the acetabular cartilage: femoroacetabular impingement as a cause of early osteoarthritis of the hip. *J Bone Joint Surg Br*. 2005;87(7):1012-8.
35. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, Brown C, Cooke TD, Daniel W, Feldman D. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum*. 1991; 34(5) 505-14.
36. Bird PA, Oakley SP, Shnier R, Kirkham BW. Prospective evaluation of magnetic resonance imaging and physical examination findings in patients with greater trochanteric pain syndrome. *Arthritis Rheum*. 2001; 44(9):2138-45.
37. Schneiders AG, Davidsson A, Horman E, Sullivan SJ. Functional movement screen normative values in a young, active population. *Int J Sports Phys Ther*. 2011;6(2):75-82.
38. Kiesel K, Plisky PJ, Voight ML. Can serious injury in professional football be predicted by a preseason functional movement screen? *N Am J Sports Phys Ther*. 2007;2(3):147-158.
39. Noyes FR, Barber SD, Mangine RE. Abnormal lower limb symmetry determined by function hop tests after anterior cruciate ligament rupture. *Am J Sports Med*. 1991;19(5):513-518.
40. Petschnig R, Baron R, Albrecht M. The relationship between isokinetic quadriceps strength test and hop tests for distance and one-legged vertical jump test following anterior cruciate ligament reconstruction. *J Orthop Sports Phys Ther*. 1998;28(1):23-31.
41. Reid A, Birmingham TB, Stratford PW, Alcock GK, Giffin JR. Hop testing provides a reliable and valid outcome measure during rehabilitation after anterior cruciate ligament reconstruction. *Phys Ther*. 2007;87(3):337-349. doi: 10.2522/ptj.20060143.
42. Barber SD, Noyes FR, Mangine RE, McCloskey JW, Hartman W. Quantitative assessment of functional limitations in normal and anterior cruciate ligament-deficient knees. *Clin Orthop Relat Res*. 1990;(255):204-214.
43. Munro AG, Herrington LC. Between-session reliability of four hop tests and the agility T-test. *J Strength Cond Res*. 2011;25(5):1470-1477. doi: 10.1519/JSC.0b013e3181d83335.
44. Hickey KC, Quatman CE, Myer GD, Ford KR, Brosky JA, Hewett TE. Methodological report: Dynamic field tests used in an NFL combine setting to identify lower-extremity functional asymmetries. *J Strength Cond Res*. 2009;23(9):2500-2506. doi: 10.1519/JSC.0b013e3181b1f77b.
45. Veale JP, Pearce AJ, Carlson JS. Reliability and validity of a reactive agility test for Australian football. *Int J Sports Physiol Perform*. 2010;5(2):239-248.
46. Youdas JW, Mraz ST, Norstad BJ, Schinke JJ, Hollman JH. Determining meaningful changes in pelvic-on-femoral position during the Trendelenburg test. *J Sport Rehabil*. 2007;16(4):326-35. PubMed PMID: 18246899.
47. Plisky PJ, Rauh MJ, Kaminski TW, Underwood FB. Star Excursion Balance Test as a predictor of lower extremity injury in high school basketball players. *J Orthop Sports Phys Ther*. 2006;36(12):911-9. PubMed PMID: 17193868.
48. Clark RC, Saxion CE, Cameron KL, Gerber JP. Associations between three clinical assessment tools for postural stability. *N Am J Sports Phys Ther*. 2010;5(3):122-30. PubMed PMID: 21589668.
-

-
49. Finnoff JT, Peterson VJ, Hollman JH, Smith J. Intrarater and interrater reliability of the Balance Error Scoring System (BESS). *PM R*. 2009; 1(1):50-4. PubMed PMID: 19627872.
 50. Bressel E, Yonker J, Kras J, Heath E. Comparison of static and dynamic balance in female collegiate soccer, basketball, and gymnastics athletes. *J Athl Train*. 2007;42:42-46.
 51. Onate JA, Beck BC, Van Lunen BL. On-field testing environment and balance error scoring system performance during preseason screening of healthy collegiate baseball players. *J Athl Train*. 2007;42(4):446-51. PubMed PMID: 18174931.
 52. Munro AG, Herrington LC. Between-session reliability of the star excursion balance test. *Phys Ther Sport*. 2010 Nov;11(4):128-32. PMID: 21055706.
 53. Hertel J, Miller S, Denegar C. Intratester and intertester reliability during the Star Excursion Balance Tests. *J Sport Rehabil*. 2000;9:104-116.
 54. Bolgia LA, Keskula DR. Reliability of lower extremity functional performance tests. *J Orthop Sports Phys Ther*. 1997;26(3):138-42. PubMed PMID: 9276854.
 55. Kea J, Kramer J, Forwell L, Birmingham T. Hip abduction-adduction strength and one-leg hop tests: test-retest reliability and relationship to function in elite ice hockey players. *J Orthop Sports Phys Ther*. 2001 Aug;31(8):446-55. PubMed PMID: 11508614.
 56. Caffrey E, Docherty CL, Schrader J, Klossner J. The ability of 4 single-limb hopping tests to detect functional performance deficits in individuals with functional ankle instability. *J Orthop Sports Phys Ther*. 2009;39(11):799-806. PubMed PMID: 19881005.